



The NASA Taxonomy What Is It and How Do We Use It?

Training for JPL Publishers

Jayne Dutra

NASA Office of the CIO

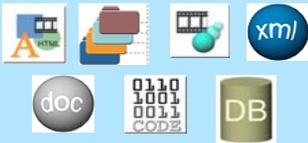
May 13th, 2004



Purpose of the NASA Taxonomy

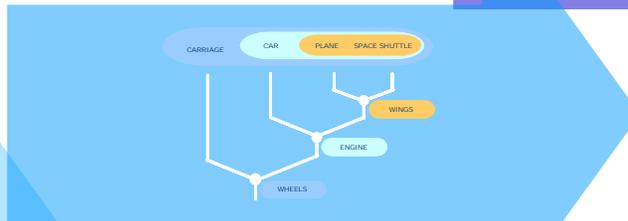


Create



- Content
- Assets

Classify



- Metadata
- Taxonomy

Discover



- Site Maps
- Search Engine
- NASA Portals
- Content Integration Networks

Finding the right information at the right time to solve the problem at hand



Project Benefits: Enable Knowledge Discovery



- Make it easy for various audiences to find relevant information quickly
 - Provide easy access for NASA resources found on the Web
 - Share knowledge by enabling users to easily find links to databases and tools as well as text
 - Deep Web and Semantic Web
 - Provide search results targeted to user interests
 - Enable the ability to move content through the enterprise to where it is needed most
- Comply with E-Government Act of 2002
- Contribute to NASA Enterprise Architecture planning



- Design process that:
 - Incorporates existing federal and industry terminology standards like NASA AFS, NASA CMS, FEA BRM, NAICS, and IEEE LOM
 - XML schema will be registered in the NASA XML namespace registry
 - Complies with metadata standards like Z39.19, ISO 2709, and Dublin Core
- Practices increase interoperability and extensibility



What is the NASA Taxonomy?

- The classification scheme is meant to encompass all of NASA web content (NASA web space) including internal as well as external material. It is a means for tagging content so it can be used and reused in different contexts.

How to Use the NASA Taxonomy

- This is a generic taxonomy from which specializations can be derived for specific purposes.
 - **Not all facets (top level categories) need to be used in each instance**
 - **Facets are repeatable**
 - **The taxonomy is modular and dynamic**
 - **Tag only as appropriate**



NATIONAL AERONAUTICS
AND SPACE ADMINISTRATION

+ Visit NASA.gov
+ Contact the NASA Curator

- TAXONOMY TOP LEVEL FACETS

+ FAQs

+ NASA METADATA

+ NASA XML

NASA Taxonomy - Top Level Facets

[Tips on using the NASA Taxonomy](#)

- [Access Security Requirements](#)
- [Audiences](#)
- [Business Purpose](#)
- [Competencies](#)
- [Content Types](#)
- [Industries](#)
- [Instruments](#)
- [Locations](#)
- [Missions and Projects](#)
- [Organizations](#)
- [Subject Categories](#)

What is the NASA taxonomy?

The NASA taxonomy is a controlled vocabulary that is designed to populate the [NASA metadata core specification](#).

It is also a means of tagging NASA content so that it can be used and reused in many different contexts.

© 2002-2004 California Institute of Technology, and Taxonomy Strategies. All rights reserved.



- + Freedom of Information Act
- + The President's Management Agenda
- + FY2002 Agency Performance Report
- + NASA Privacy Statement, Disclaimer, and Accessibility Certification
- + Freedom to Manage



Contact the NASA Curator
NASA Official: Jayne Dutra
Last Updated: March 11, 2004
+ Contact NASA



Using the Taxonomy



- NASA Taxonomy provides controlled vocabularies used to populate elements of more complex metadata schema such as the Dublin Core (www.dublincore.org)
- The taxonomy facets map to the DC and NASA metadata tags

Field	Name space	Definition	Data Type or Source
Type	dc:type	The nature or genre of the content of the resource	Values come from NASA taxonomy facet: Content Types



The Taxonomy as a Resource

- Consistent content tagging
 - Document, Project, Content, Records and Digital Asset Management
 - Project libraries and CM systems for external sites
 - Agency Website Registration System
 - UDDI
 - RSS content syndication
- Facilitate re-using content on public and internal portals
- Improve search & navigation
 - Verity K2, and other search engines
 - *Facet search and navigation applications*



Tagging Strategies



- Tag your home page first
 - Think about your most significant material and give it priority
- Tag when you redesign your site
- Embed tags in any automated systems you may be using



Attribute	Description
Content Type	Types of content by genre and function.
Audiences	Audiences for a content item.
Organizations	The NASA organization, affiliates, and business partners.
Missions and Projects	NASA's business lines.
Industries	Industries NASA does business with.
Locations	Locations on and off Earth.
Business Purpose	Purpose of business records.
Competencies	Technical specialties.



NASA Taxonomy – Dublin Core Map Fragment



Dublin Core Elements	Definition	NASA Taxonomy Mapping
Creator Creator Affiliation	Content maker Organization employing creator	dc:creator dc:creator.affiliation
Subject	Content topic	dc.subject
Publisher	Publisher of this manifestation	dc:publisher
Format	File format of the resource	dc:format
Type	Content Genre	dc:type
Coverage	Space, period, date, jurisdiction, etc.	dc:coverage.locations dc:coverage.chronology
Audience	Content audience	dcTERM:audience
Non DC	NASA missions and projects	nasa:missionsProjects
Non DC	Business purpose	nasa:businessPurpose
Non DC	Technical competencies	nasa:competencies
Non DC	Standard industry categories	naics:industries



Indexing rules...

Simplified for publisher indexing



Rule	Description
Specificity rule	Apply the most specific terms when tagging assets. Specific terms can always be generalized, but generic terms cannot be specialized.
Repeatable rule	All attributes should be repeatable. Use as many terms as necessary to describe <i>What the asset is about</i> and <i>Why it is important</i> . Storage is cheap. Re-creating content is expensive.
Appropriateness rule	Not all attributes apply to all assets. Only supply values for attributes that make sense.
Usability rule	Anticipate <i>how the asset will be searched for</i> in the future, and <i>how to make it easy to find it</i> . Remember that search engines can only operate on explicit information.



NASA Taxonomy – Example



Metadata Syntax Example

<meta name = “dc.title” content=“Jet Propulsion Laboratory”>

<meta name = “dc.publisher” content=“Jet Propulsion Laboratory”>

<meta name = “dc.date” content=“2004-05-07”>

<meta name = “dc.type” content=“web pages”>

<meta name = “dc.subject” content=“Lunar and Planetary exploration”>

<meta name = “dcterms.audience” content=“External”>

<meta name = “nasa:missionsProjects” content=“Space Sciences, Planetary Missions”>

<meta name = “nasa:competencies” content=“Public Communications and Outreach”>



Tagging Example- JPL Homepage

Element	Value
Title	Jet Propulsion Laboratory
Description	The Jet Propulsion Laboratory is the lead U.S. center for robotic exploration of the solar system.
Date	2004-05-07
Format	Text/html
Keywords	Caltech; California Institute of Technology; www.jpl.nasa.gov
Publisher	Jet Propulsion Laboratory
Type	Web Pages
Subject	Lunar and planetary exploration
Coverage	Solar System
Access Rights	Public Access Information
Audience	External
Competencies	Public Communications and Outreach
Missions and Projects	Space Sciences; Planetary Missions
Business Purpose	Administrative



Tagging Example



Mission
Homepage

Example:
Topex
Poseidon
satellite mission



<http://topex-www.jpl.nasa.gov>



Tagging Example- Mission Homepage

Element	Value
Title	Ocean Surface Topography
Description	Ocean Surface Topography seeks to study the sea surface height. Maintaining a database of ocean surface topography can help predict short-term changes in weather and longer-term patterns of climate.
Date	2004-04-03
Format	text/html
Keywords	Ocean Topography; Sea Level; SSH; TOPEX; Poseidon; Jason; Jason1; sea surface height
Publisher	Jet Propulsion Laboratory
Type	Web Pages
Subject	Geosciences
Coverage	Earth
Access Rights	Public Access Information
Audience	External
Competencies	Oceanographic science; Public communications and outreach
Missions and Projects	Space Sciences; Planetary Missions
Business Purpose	Administrative



Tagging Example- Science Page

Element	Value
Title	EOSDIS Page
Description	Portal to the Earth Observing System Data and Information System (EOSDIS) including links to the Earth Observatory, Missions information, Data Gateway, Data Products, DAACs and other cooperating centers.
Date	2004-04-12
Format	text/html
Keywords	Distributed Active Archive Centers; DAAC Alliance
Publisher	Earth Sciences Directorate; Goddard Space Flight Center
Type	Web Pages
Subject	Geosciences
Keywords	Distributed Active Archive Centers; DAAC Alliance
Access Rights	Public Access Information
Audience	External
Competencies	Scientific; Computer and Information Technologies
Missions and Projects	Earth Observing Data and Information System



Discussion



- Questions?
- Feel free to call or e-mail me!
 - ext. 4-6948
 - Jayne.E.Dutra@jpl.nasa.gov